



PCT/GB 2004 / 001354



INVESTOR IN PEOPLE

## PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

The Patent Office  
Concept House  
Cardiff Road

Newport  
South Wales  
NP10 8QQ

04 MAY 2004

WIPO

PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 27 April 2004

THE PATENT OFFICE	
For Official use only	RIM
18 NOV 2003	
RECEIVED BY FAX	
Your reference VXT (UK)	T

<b>The Patent Office</b>	<b>Request for grant of a Patent</b>
	<b>Form 1/77 Patents Act 1977</b>
1. Title of invention	
<b>COMPUTER BASED SYSTEM FOR DIGITAL MEDIA PREPARATION</b>	
2. Applicant's details	
<input checked="" type="checkbox"/> 18 NOV 2003	0326796.0
<input type="checkbox"/> 2a First or only applicant	
If applying as a corporate body: Corporate Name	
<b>Internet Pro Video Limited</b>	
Country	
Great Britain	
2b If applying as an individual or partnership	
Surname	
Forenames	
2c Address	
Mount Pleasant House 2 Mount Pleasant Huntingdon Road Cambridge	
UK Postcode CB3 0RN	
Country Great Britain	
ADP Number	
<b>0860 495 1001</b>	

**4 Reference Number**

**VXT (UK)**

**5 Claiming an earlier application date**

An earlier filing date is claimed:

Yes  No

Number of earlier  
application or patent number

Filing date

15 (4) (Divisional)    8(3)    12(6)    57(4)

**6 Declaration of priority**

Country of filing	Priority Application Number	Filing Date
GB	0307884.7	7.4.2003 (7 April 2003)

**7 Inventorship**

The applicant(s) are the sole inventors/joint inventors

Yes

No

**8 Checklist**

Claims 3 Continuation sheets

Description 9

Abstract 0

Drawings 8

*only fm*

Priority Documents Yes/No

Translations of Priority Documents Yes/No

Patents Form 7/77 Yes/No

Patents Form 9/77 Yes/No

Patents Form 10/77 Yes/No

**9 Request**

We request the grant of a patent on the basis  
of this application

Signed: *Origin Limited* Date: *18 November 2003*  
(Origin Limited)

# DUPLICATE

1

Computer based system for digital media preparation.

## Technical Field

This invention relates to a computer software system for the preparation of digital media.

## Background Art

5 Application software for editing digital video is an extremely sophisticated and powerful tool because it is primarily designed for, and sold to, the video professional. Such an individual requires access to many complex functions and is prepared to invest time and effort in learning to become skilled in their use. Historically, the terminology and conventions of Digital Editing have evolved from a traditional film editing environment where rushes are cut and spliced together to tell a story or follow a script. As digital mixer technology advanced new techniques were combined with these conventional methods to form the early pioneering software based digital editors.

To the video or film professional editing is second nature and the complexities of a time-based media go unnoticed since, having already grasped concepts and learned processes, they 15 are able to concentrate on the nuances of different editing packages, of which there are many.

Conventionally these packages, through the use of a Graphical User Interface (GUI), attempt 20 to provide an abstraction of the media in terms of many separate tracks of video and audio. These are represented on the output device in symbolic fashion and provision is made for interacting with these representations using an input device such as a mouse. Typically the purpose is to create a new piece of media as an output file, composed by assembling clips or segments of video and audio along a timeline that represents the temporal ordering of frames. Special effects such as wipes and fades can be incorporated, transparent overlays can be added, colour and contrast can be adjusted. The list of manipulations made possible by 25 such tools is very long indeed. A typical system is described in, for example, Foreman; Kevin J., et al, "Graphical user interface for a video editing system", U.S. Patent. 6,469,711.

It is possible, however, that an individual who is a consumer of media, rather than a producer, may need to perform a simple editing operation on a media file in order to accomplish their primary task; for example to give a multi-media presentation. In this case 30 such tools have their drawbacks. They may be too expensive to justify individually, or to

have enough of in order to be available when or where needed. The limited amount of use and the small fraction of the capabilities used in such situations may make them uneconomic. The steep learning curve associated with such tools may mean that an inappropriate amount of effort is expended on something that is not the primary occupation or concern of the tool user. For occasional or infrequent use there will be reluctance on the part of any user repeatedly to switch environments or learn and relearn new tools to perform simple last minute tasks.

Work has been carried out with the view of improving the interaction between a user and a video editor by providing 'intelligent' operations; The 'Silver' project (Juan P. Casarcas, "SILVER: An Intelligent Video Editor," ACM CHI'2001 Student Posters, Seattle, WA, March 31-April 5, 2001, pp. 425-426) uses 'smart selection' to assist the user to find 'in' and 'out' points. The 'in' and 'out' points are roughly set by the user and then 'snap' to a boundary, which could be a shot change or the silence between spoken words, or other similar features. Video and audio boundaries typically will not line up so the system provides some 'fixing-up' functions to smooth the edit boundary.

Conventionally, video editors are application programs that run on high-end PCs and workstations, under desktop-oriented operating systems such as Microsoft Windows or Apple's Mac OSX, often with high-resolution screens and high-bandwidth network connectivity. The viewing of media files, however, can take place on an ever-expanding list of devices with many different capabilities, such as laptops, mobile PDAs with wireless connectivity, mobile phones, set-top boxes and hard-disc based personal video recorders (PVRs). The concept of a simple media manipulation tool integrated into the media player component is as relevant in these cases as it is in that of the standard PC, possibly more so since, for example, a PVR may not have a run-time environment capable of running external applications such as video editors.

Another class of device that is becoming ever more capable of media manipulation is the mobile phone. Such devices now have the ability to capture, display and transmit moving images, but, conventionally, are not thought of as a platform for editing video. There is no reason, however, why simple editing operations should not be applied here in order to enhance even the simplest and shortest of video presentations. Mobile phones present a unique set of challenges to the user interface component of any application. First and foremost the display area is extremely limited and so immediately rules out multi-level menus, timelines and story-boards. Secondly, the user interface is extremely constrained:

there is no mouse input, only a few options can be displayed at a time, and all interaction must be performed using a set of navigation buttons (which may vary in position and size according to the hardware manufacturer). Thirdly, the user expects to be able to perform any action one-handed.

- 5 Accordingly, these are the attributes of a media preparation tool that is appropriate to the needs of such a device.
  - Simple and intuitive to use; in particular, little time and effort is required to learn enough to accomplish the task in hand.
  - Efficient use of screen area; no menus, timelines or story-boards.
- 10 • Efficient use of user input interface.
  - Efficient editing model that allows simple trimming operations to be performed simply, whilst permitting more complex tasks to be carried out.
  - Predicts, where possible, the preferences of the user as regards editing limits.

#### **Disclosure of Invention**

- 15 The invention relates to a method called VXT for simple video message preparation, analogous to the predictive text editing for mobile 'TXT'ing. VXT does not use the conventional editing semantics of 'in' and 'out' points; it determines edit limits using rules that are updated through user feedback and minimises the typical number of user interactions required to perform a simple video 'trimming' task.
- 20 Briefly, the invention works as follows.

According to one aspect of the invention a Graphical User Interface (GUI) input interface for editing is defined.

In the preferred embodiment of this aspect of the invention the controls consist of five buttons:

- 25
  - one for video 'forward' shuttle;
  - one for video 'backward' shuttle;
  - one button meaning 'include';
  - one button meaning 'exclude';
  - one button meaning 'apply'.

According to another aspect of the invention a Graphical User Interface (GUI) output interface for editing is defined for feedback to the user.

In the preferred embodiment of this aspect of the invention the graphical elements consist  
5 of

- an 'edit bar' graphic on the display; consisting of a coloured rectangular area.
- a 'frame pointer' that marks the current frame on the edit bar.
- an 'include' graphic which overlays the corresponding frame and consists of a green 'tick';
- an 'exclude' graphic which overlays the corresponding frame and consists of a red 'cross'.

10

According to another aspect of the invention the sequence of actions from the user loading a piece of digital media to the user applying the edits is called a '*session*'; the first operation the user performs during a session is called the '*initial selection*'; subsequent operations that the user performs are called the '*refinement phase*'; a frame or frames that are in the final edit are  
15 'included'; those that are not are '*excluded*'; an operation that causes a number of frames to change state from '*excluded*' to '*included*' or vice-versa is called a '*grow*' operation; the actual number of frames that change state from '*excluded*' to '*included*', or vice-versa, during a grow operation is called the '*support*'.

15

According to another aspect of the invention means are provided for storing as variables in a  
20 computer memory information about the history of interactions between the user and the video preparation tool; these are called '*session variables*' and assist the user to determine the limits of initial selection.

20

In the preferred embodiment of this aspect of the invention an integer variable used for prediction called  $p$  is used automatically to determine the number of frames labelled as  
25 'included', as a proportion of the initial length of the clip, when the user makes the initial selection. When the program is used for the first time ever this variable is set to an arbitrary initial value, for example, 4. If the length of the clip in frames is  $L$  then the support is given by  $s = L/p$ . For example, if  $s$  equals 4 and  $L$  equals 100 then the support  $s$  equals 25 frames. Therefore, if the user nominates a particular frame as being 'included' then the system  
30 determines that 25 frames previous, and 25 frames subsequent, to this frame, may also be included.

According to another aspect of the invention means are provided for using and updating the 'session variables' to assist the user to determine the limits of editing operations that occur during the refinement phase.

- In the preferred embodiment of this aspect of the invention, after an editing session is complete the actual number of frames ( $f$ ) included in the final video message is read and is used to derive a new value of  $p$  as follows:  $p(\text{new}) = 2L/f$ . So, for example, if the length of the final message is 40 frames then the new value of  $p$  reflects the fact that fewer frames were actually required than were predicted, and the predicted  $p$  for the next edit session becomes  $200/40 = 5$ . Assuming an initial length of 100 frames in the next editing session, a support value  $s$  equal to 20 frames is used.

According to another aspect of the invention means are provided for storing as variables in a computer memory information about the history of interactions between the user and the video preparation tool; these are called '*session variables*' and assist the user to determine the limits of edit operations during the refinement phase.

- In the preferred embodiment of this aspect of the invention a vector of integer variables  $r(i)$  is used to model how the user refines the initial edit; the value of  $r(i)$  is equal to the support in frames for the  $i^{\text{th}}$  refinement edit and is used to determine the number of frames to be labelled as 'included' during refinement phases.

According to another aspect of the invention means are provided for using and updating the 'session variables' to assist the user to determine the limits of editing operations that occur during the refinement phase.

- In the preferred embodiment of this aspect of the invention, any operation that results in a change of state of a frame from 'excluded' to 'included' is treated as a new edit and causes the index  $i$  in  $r(i)$  to increment. The value of  $r(i)$  is equal to the support obtained during a 'grow' operation and is preset to some arbitrary value the very first time the program is used. For operation  $i$ , such a 'grow' operation changes the state of  $r(i)$  frames to 'included' and any further editing within the limits of these frames causes the value of  $r(i)$  to be updated accordingly.

According to another aspect of the invention means are provided for the user to select the region of the video message that is of interest.

In the preferred embodiment of this aspect of the invention the user operates the 'forward' and 'backward' shuttle buttons to find a representative frame in the part of the clip that is 'of most interest'. The desired frame is displayed along with smaller, undersampled versions of the previous and following frames.

- 5 According to another aspect of the invention means are provided to feedback to the user, without the user having to preview the edit, frames that are 'included' and 'excluded'.

In the preferred embodiment of this aspect of the invention the 'edit bar' represents the video clip being edited and a pointer in the 'edit bar' indicates the frame currently being viewed. Regions of the bar that are green represent 'included' sections; regions that are red represent 'excluded' sections. Prior to any editing taking place the bar is completely red, meaning that all the frames are 'excluded'.

According to another aspect of the invention means are provided to feedback to the user, involving the user previewing the edit, frames that are 'included' and 'excluded'.

15 In the preferred embodiment of this aspect of the invention each frame that is 'included' is overlaid with a green 'tick' and each frame that is 'excluded' is overlaid with a red cross. The user can review these frames using the forward and backward shuttle controls.

According to another aspect of the invention means are provided for the user to manipulate the region of the video message that is included.

20 In the preferred embodiment of this aspect of the invention the user operates the 'include' button to grow regions of the video clip for inclusion in the final edit. Assuming that the user has stopped at a frame in a region of interest the interaction is as follows:

- If the 'include' button is pressed once the part of the edit bar under the frame pointer goes green to indicate that only the current frame is included; the rest of the bar remains unchanged.
- 25 • If the 'include' button is pressed once more, a region corresponding to the support before and after the frame pointer position goes green to indicate that this region is included; the rest of the bar remains unchanged.
- If the 'include' button is pressed once more, a region from the start of the bar up to the pointer and a region corresponding to the support after the frame pointer position goes green to indicate that all the frames from the beginning of the video to the current

position, are included, and a number of frames after the current position corresponding to the support are also included.

- If the 'include' button is pressed once more, a region from the end of the bar back to the pointer and a region corresponding to the support before the frame pointer position goes green to indicate that all the frames from the current position to the end of the video are included, and a number of frames before the current position corresponding to the support are also included.
- Further presses repeatedly cycle round the four above cases.

- In another embodiment of this aspect of the invention the user operates two 'handles' on the edit bar that define the start and end of the included region, respectively

The user also operates the 'exclude' button to grow regions of the video clip for exclusion from the final edit. Assuming that the user has stopped at a frame in a region of interest the interaction is as follows:

- If the 'exclude' button is pressed once the part of the edit bar under the frame pointer goes red to indicate that only the current frame is 'excluded'; the rest of the bar remains unchanged.
- If the 'exclude' button is pressed once more, a region corresponding to the support before and after the frame pointer position goes red to indicate that this region is 'excluded'; the rest of the bar remains unchanged.
- If the 'exclude' button is pressed once more, a region from the start of the bar up to the pointer goes red to indicate that all the frames from the start of the video to the current position are 'excluded'.
- If the 'exclude' button is pressed once more, a region from the pointer to the end of the bar goes red to indicate that all the frames from to the current position to the end of the video are excluded.

Further presses repeatedly cycle round the four above cases.

According to another aspect of the invention means are provided for the user to export the edited video message.

In the preferred embodiment of this aspect of the invention the user operates the 'apply' button to export the edited video message.

According to another aspect of the invention means are provided for user to select further options prior to completion:

- 5 In the preferred embodiment of this aspect of the invention the user selects, through interaction with a menu, the following:
  - \* add 'fades' where frames have been deleted.
  - \* add 'transitions' where frames have been deleted.
  - \* add a background music track
- 10 \* add text annotation.

According to another aspect of the invention if any editing operation results in a single stationary frame being displayed to the user then this frame can be treated as a still image and processed separately.

- 15 In the preferred embodiment of this aspect of the invention the system monitors the support for the currently displayed frame and, if this is equal to one, asks the user via a message box whether this frame is required as a still; if the user replies 'yes' then the still is captured and stored, and the editing session can then proceed.

#### Industrial Applicability

As a simple example of the use of the invention consider this scenario. Using a built-in camera a user of a mobile phone captures a short segment of video from a birthday party and wishes to trim the segment. This trimming operation is wanted in order, both to focus in on the moment when the children blow out the candles on the birthday cake, and to minimise the cost of mailing the video segment to friends and family. The video segment is shuttled until the actual frame when the candles go out is displayed. The "include" button is pressed twice and the preparation tool, based on the past history of user interaction, determines that three seconds of video before and after the chosen frame should be included in the edit. The user runs to the start of the 'included' region and, using the 'include' button, adds more frames to the final edit. The user then quickly runs forward and backward checking that green 'tick' markers appear in the part of the clip of interest; then the 'apply'

button is pressed and the editing process is completed. The system measures the actual number of frames set as 'included' and updates the memory variables used for prediction.

**Claims.**

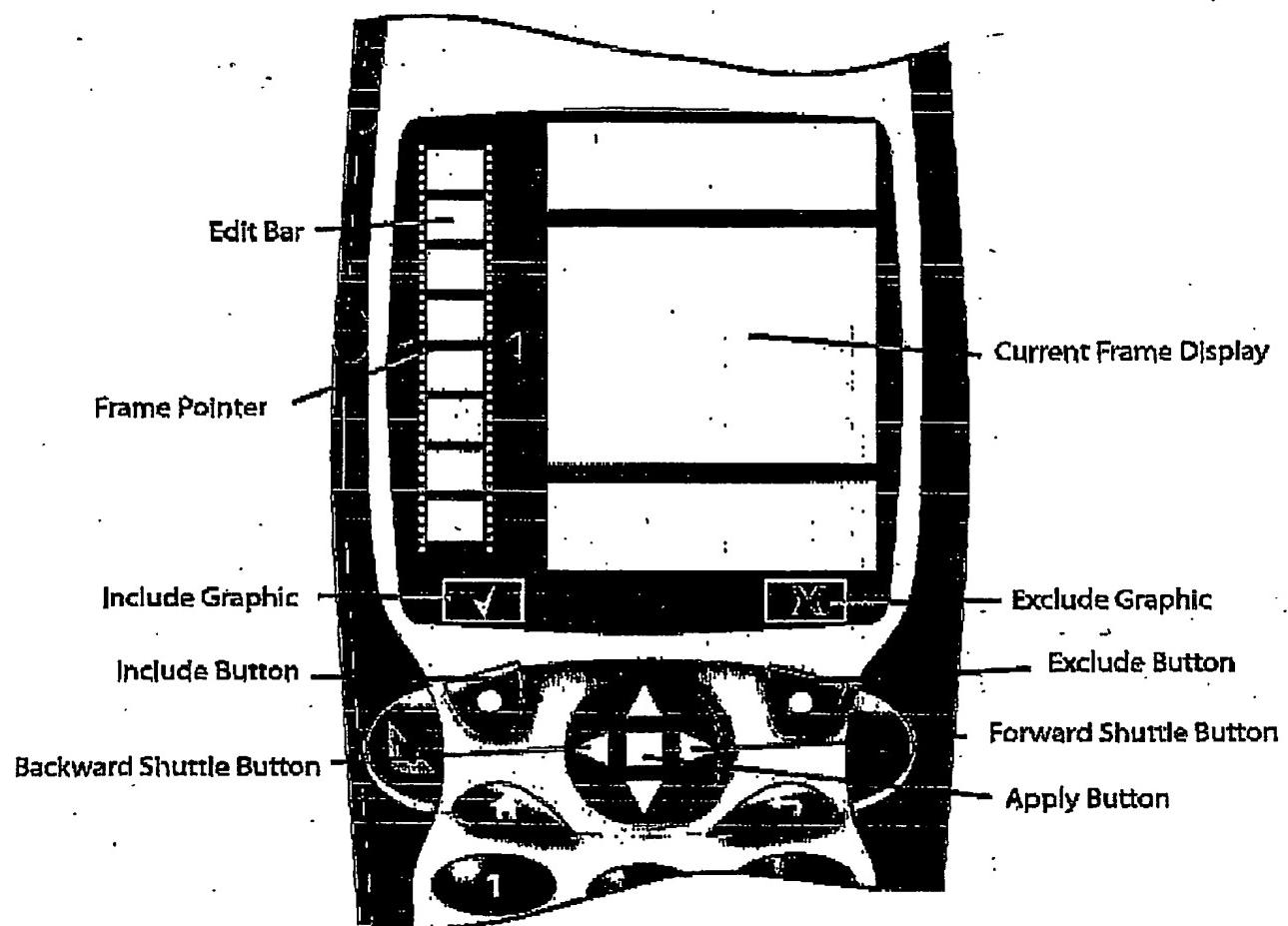
What is claimed is:

1. A computer based system for preparing digital media that adds the capability for media-preparation to media players wherein information held in computer memory describes the interaction between the system and the user of the system and is retained between uses of the preparation system, which information is used to predict the preferences of the user during subsequent uses of the preparation system, and which information additionally is used to assist the user in performing a media preparation task, and which information additionally is updated as a result of the history of uses of the preparation system.  
10
  2. The digital media preparation system of claim 1 comprising:
    - Input interfaces for forward and backward video transport;
    - Input interfaces for including and excluding video frames;
    - Input interfaces for applying the edit operation;  
15
  - A graphical output interface that schematically represents the digital media in terms of regions representing frames to be included in, and regions representing frames that are to be excluded from, the final edit.
  - A 'frame pointer' that marks the current frame on the schematic representation.
  - An 'include' graphic on the display that overlays a frame and which confirms that that frame is to be retained.  
20
  - An 'exclude' graphic on the display that overlays a frame and which confirms that that frame is to be discarded.
  - One or more variables held in computer memory that hold information describing the interaction between the system and the user of the system.
- 25     3. The digital media preparation system of claims 1 & 2 wherein the sizes and positions of the regions representing included frames are modified using the input interface for including frames.
- 30     4. The digital media preparation system of claims 1 & 2 wherein the sizes and positions of the regions representing excluded frames are modified using the input interface for excluding frames.

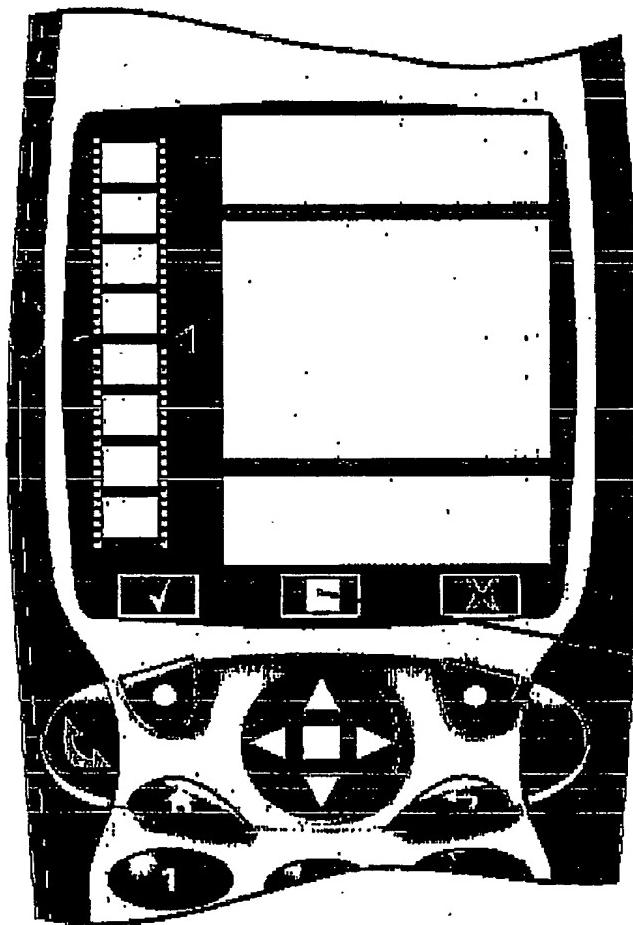
5. The digital media preparation system of claims 1 - 4 wherein the input interfaces for including and excluding frames, forward and backward transport, and applying the edit, are implemented as push buttons.
6. The digital media preparation system of claims 1 - 4 wherein the input interfaces for including and excluding frames, forward and backward transport, and applying the edit, are implemented as graphics on a display.
7. The digital media preparation system of claims 1 - 6 wherein the sizes and positions of the regions to be retained and discarded are controlled using multiple button presses.
8. The digital media preparation system of claims 1 - 7 wherein one button press selects one frame only, that being the currently viewed frame, a second press selects a region centred on the current frame, a third press selects the region from the beginning of the digital media up to the current frame and a fourth press selects the region from current frame to the end of the digital media, the sequence then continuing to cycle as further button presses are made.
- 15 9. The digital media preparation system of claims 1 - 8 wherein the action of selecting a single frame nominates the selected frame as a still image which may then be subjected to further processing appropriate to the properties of still images.
- 10 10. The digital media preparation system of claims 1 - 8 where media includes, but is not confined to, video and audio.
- 20 11. The digital media preparation system of claims 1 - 8 where preparation includes, but is not confined to, the operations of one or more of:  
Editing; trimming; annotating; effects; transitions; appearance; presentation;
12. The digital media preparation system of claim 1 wherein the information held in computer memory determines the number of frames to set, in response to user interaction, as being included or excluded.
- 25 13. The digital media preparation system of claim 12 wherein the information held in computer memory determines the number of frames to set, in response to the first user interaction of a session, as being included.

14. The digital media preparation system of claim 12 & 13 wherein the information held in computer memory determines the number of frames to set, in response to user interactions other than the first of a session, as being included or excluded.

15. The digital media preparation system of claim 12 - 14 wherein the information held in computer memory is updated as a result of user interaction.  
5

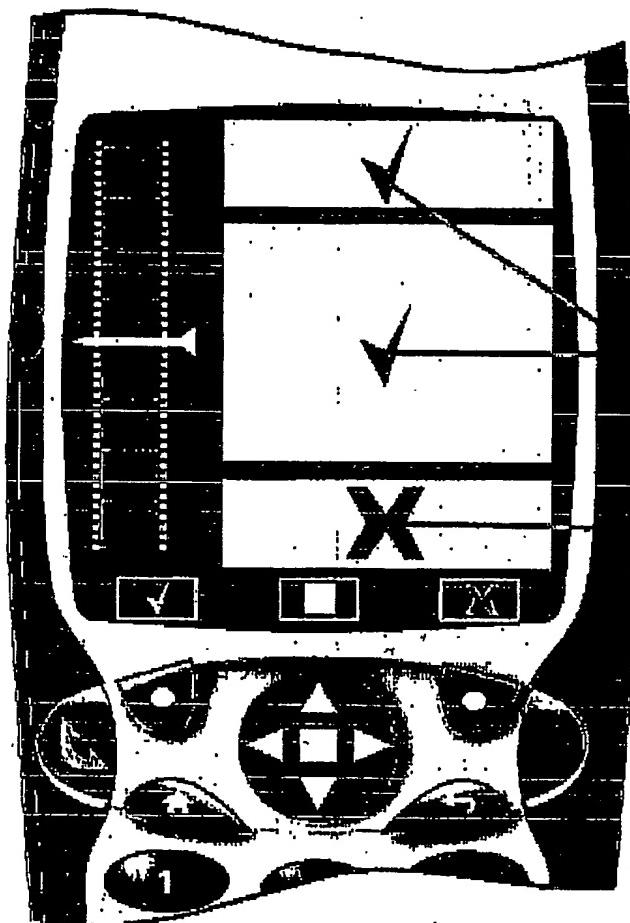


2/8



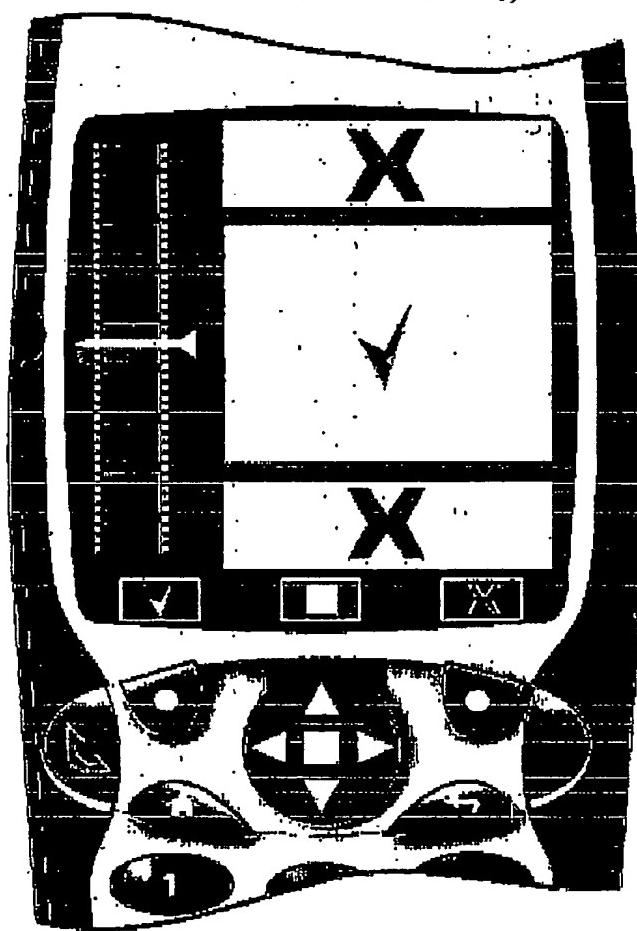
3/8

GEOMETRY TEST

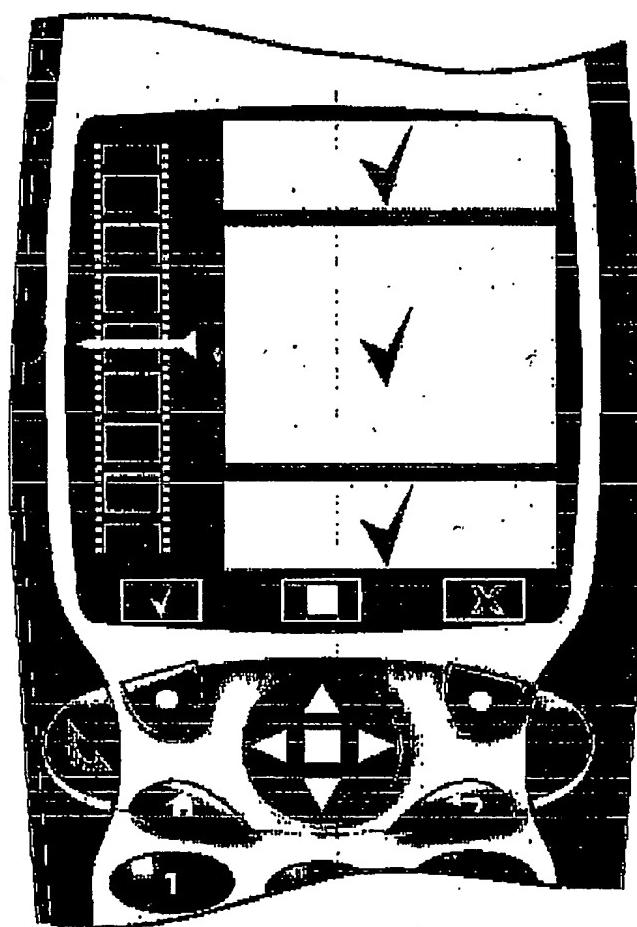


Include Graphic  
(Superimposed on Frame)

Exclude Graphic  
(Superimposed on Frame)

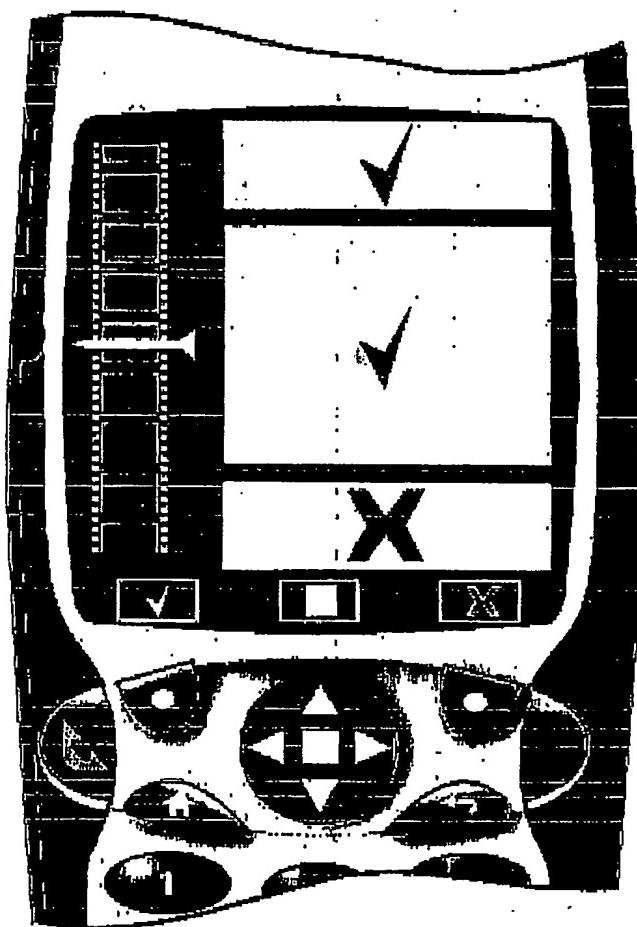


5/8

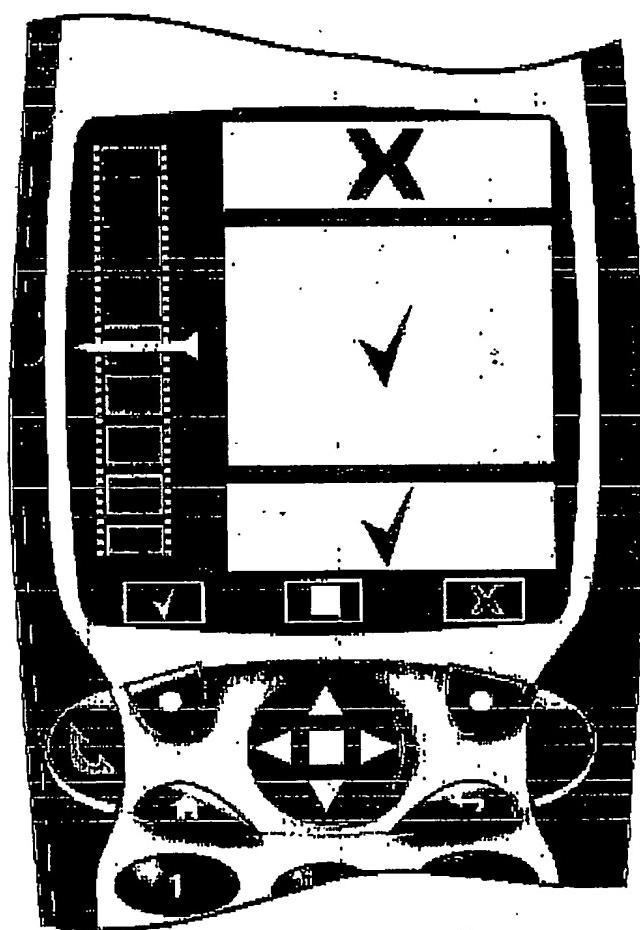


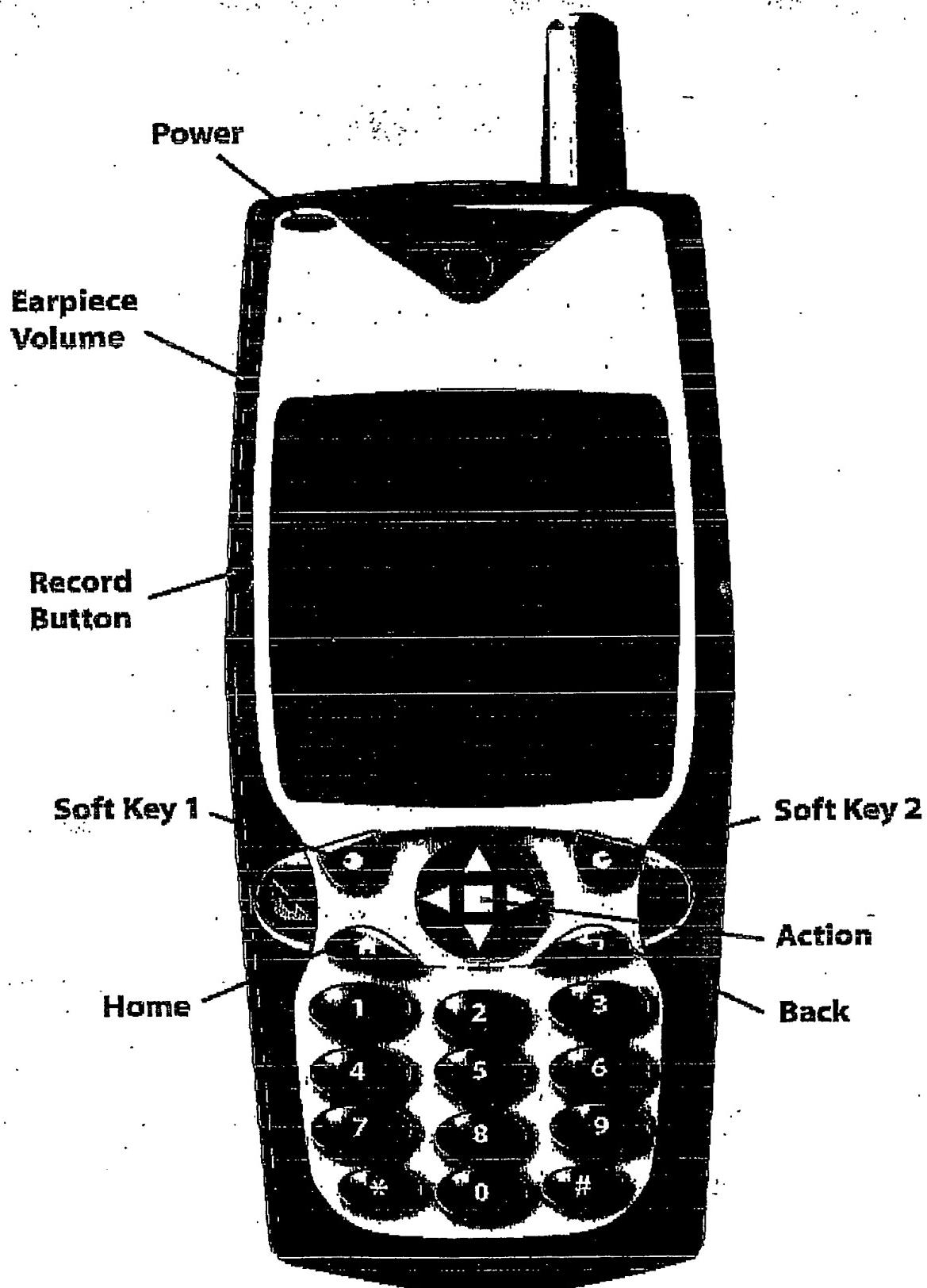
SEARCHED INDEXED SERIALIZED FILED

6/8

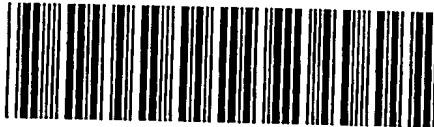


7/8





PCT/GB2004/001354



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT OR DRAWING
- BLURRED OR ILLEGIBLE TEXT OR DRAWING
- SKEWED/SLANTED IMAGES
- COLOR OR BLACK AND WHITE PHOTOGRAPHS
- GRAY SCALE DOCUMENTS
- LINES OR MARKS ON ORIGINAL DOCUMENT
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.